# **Trích xuất nhanh sóng hô hấp từ quang thể tích đồ: Phương pháp tiếp cận bằng corr-encoder sâu**

Nhiều thông tin liên quan đến hô hấp được chứa trong tín hiệu quang thể tích đồ (PPG), thông qua các thay đổi trong lưu lượng máu tĩnh mạch, nhịp tim và thể tích tâm thu. Chúng tôi nhằm tận dụng thực tế này bằng cách sử dụng một khung học sâu mới dựa trên bộ tự mã hóa tích chập (convolutional autoencoder) được điều chỉnh lại mục đích. Mô hình corr-encoder của chúng tôi nhằm mã hóa tất cả thông tin hô hấp liên quan có trong dạng sóng quang thể tích đồ, và giải mã nó thành dạng sóng tương tự với tham chiếu hô hấp tiêu chuẩn vàng - Capnogram. Mô hình được áp dụng trên hai bộ dữ liệu PPG, cụ thể là Capnobase và BIDMC. Chúng tôi chứng minh rằng mô hình có khả năng tạo ra dạng sóng hô hấp tiếp cận tiêu chuẩn vàng, đồng thời tạo ra ước tính tần suất hô hấp đạt tiêu chuẩn cao nhất hiện nay. Chúng tôi cũng chỉ ra rằng khi nói đến việc nắm bắt các đặc điểm dạng sóng hô hấp nâng cao hơn như chu kỳ hoạt động, mô hình của chúng tôi phần lớn không thành công. Lý do được đề xuất cho điều này, dựa trên một nghiên cứu trước đây về PPG trong tai, là vì các biến đổi hô hấp trong PPG ở ngón tay yếu hơn nhiều so với các vị trí ghi khác. Quan trọng là, mô hình của chúng tôi có thể thực hiện các ước tính dạng sóng này trong một phần nhỏ của mili giây, cho phép nó tạo ra hơn 6 giờ dạng sóng hô hấp trong một giây. Hơn nữa, chúng tôi cố gắng giải thích hành vi của trọng số hạt nhân trong mô hình, cho thấy rằng một phần mô hình của chúng tôi trực quan chọn các tần số thở khác nhau. Mô hình được đề xuất trong công trình này có thể giúp cải thiện tính hữu ích của các thiết bị đeo dựa trên PPG cho mục đích y tế, nơi thông tin hô hấp chi tiết là cần thiết.

## **1. Giới thiệu**

Công nghệ sức khỏe đeo được hứa hẹn sẽ cách mạng hóa y học thông qua các con đường chăm sóc cá nhân hóa và theo dõi dài hạn các bệnh mãn tính. Tuy nhiên, những gì thiết bị đeo được đạt được từ sự thuận tiện và số lượng dữ liệu, chúng thường bị hạn chế từ góc độ thông tin khả dụng. Một ví dụ điển hình về điều này là việc sử dụng quang thể tích đồ (PPG) cho các ứng dụng hô hấp. Mặc dù việc sử dụng PPG cho đến nay đã chứng minh độ chính xác cao trong việc theo dõi tần suất hô hấp, nhưng nhiều thông tin hô hấp được mã hóa trong PPG vẫn chưa được khai thác.

Sự gia tăng tần suất hô hấp khi nghỉ ngơi có thể chỉ ra các rối loạn hô hấp, chẳng hạn như bệnh phổi tắc nghẽn mãn tính (COPD), các bệnh phổi hạn chế như xơ phổi và tăng áp phổi [1]. Quan trọng là, thông tin mẫu hô hấp thêm như chu kỳ hoạt động hít vào có thể giúp phân biệt giữa các loại bệnh phổi cụ thể [2]. Vì mục đích này, chúng tôi tập trung vào việc thu hẹp khoảng cách giữa quang thể tích đồ và giám sát hô hấp tiêu chuẩn vàng thông qua một khung mã hóa-giải mã tích chập đa tầng lấy PPG làm đầu vào và đầu ra là dạng sóng hô hấp chính xác cao. Chúng tôi chứng minh rằng các dạng sóng hô hấp này không chỉ hữu ích để ước tính tần suất hô hấp, mà còn có thể so sánh về sai số tuyệt đối trung bình với các phép đo hô hấp tiêu chuẩn vàng. Hơn nữa, mô hình học sâu đơn giản này có chi phí tính toán thấp, làm cho việc triển khai trong các thiết bị đeo trở nên khả thi cao.

### **1.1. Quang thể tích đồ và hô hấp**

Quang thể tích đồ (PPG) là phép đo không xâm lấn về thể tích máu. Nó hoạt động bằng cách phát ra ánh sáng qua da và vào mô bên dưới, với một hoặc nhiều đi-ốt phát quang, sau đó đo ánh sáng truyền qua mô (PPG truyền qua) hoặc phản xạ trở lại (PPG phản xạ) đến một hoặc nhiều quang đi-ốt. Máu hấp thụ ánh sáng, và do đó khi có nhiều máu hơn, ít ánh sáng được phản xạ trở lại quang đi-ốt. Vì vậy, PPG có thể đo được những thay đổi trong thể tích máu xảy ra với nhịp đập và hô hấp.

Có một gradient áp suất giữa tĩnh mạch ngoại vi và tim dẫn đến sự trở lại của máu khử oxy về tim (được gọi là trở lại tĩnh mạch). Vì tim nằm trong khoang ngực, và áp suất trong lồng ngực phải thay đổi để cho phép luồng không khí vào và ra khỏi phổi, hô hấp do đó điều chỉnh gradient áp suất này và trở lại tĩnh mạch. Khi chúng ta thở ra, áp suất trong lồng ngực tăng lên để đẩy không khí ra khỏi phổi và điều này lần lượt làm giảm gradient áp suất giữa tim và tĩnh mạch ngoại vi. Gradient áp suất nhỏ hơn có nghĩa là ít máu trở về tim hơn, và sự gia tăng thể tích máu tại vị trí cảm biến PPG. Hơn nữa, khi phổi co thắt, nhiều máu hơn chảy từ tĩnh mạch phổi vào tâm nhĩ trái, sau đó dẫn đến tăng thể tích tâm thu và do đó tăng biên độ xung tại vị trí cảm biến PPG. Biên độ xung lớn hơn khi thở ra cũng đi kèm với tốc độ xung chậm hơn. Cả ba thay đổi này đều bị đảo ngược trong quá trình hít vào [3].

Để khôi phục thông tin hô hấp từ PPG, các kỹ thuật thường dựa vào việc trích xuất một hoặc nhiều trong ba chế độ hô hấp (biến đổi cường độ, biến đổi biên độ xung và biến đổi khoảng xung). Các phương pháp phổ biến để ước tính tốc độ hô hấp bao gồm tìm đỉnh phổ tương ứng với hô hấp trong tín hiệu PPG [4] và lọc dải thông để phát hiện các đỉnh hô hấp trong miền thời gian [5].

Quang thể tích đồ cũng có thể được phân tách thành các thành phần tần số có ý nghĩa vật lý (các hàm chế độ nội tại) thông qua việc sử dụng phân tách chế độ thực nghiệm (EMD) hoặc tương đương đa biến của nó (MEMD) [6]. Điều này đã được kết hợp với các phương pháp như phân tích thành phần chính để chọn các chế độ hô hấp để đưa ra ước tính tốc độ hô hấp chính xác hơn [7]. Cũng đã được chứng minh rằng các hàm chế độ nội tại hô hấp đủ chi tiết để xác định tỷ lệ giữa thời gian hít vào và thở ra, cho phép phân loại bệnh phổi tắc nghẽn mãn tính (COPD) [8]. Các phương pháp trên thường được áp dụng cho cả ba biến đổi hô hấp, và các ước tính tốc độ hô hấp thu được sau đó có thể được kết hợp thành một ước tính tốc độ hô hấp cuối cùng chính xác hơn [9].

### **1.2. Capnography**

Capnography đề cập đến việc đo nồng độ CO2 theo thời gian trong quá trình thở [10]. Nó chủ yếu được đo từ bệnh nhân đã được đặt nội khí quản trong các kịch bản như phẫu thuật. Khi bệnh nhân hít vào, nồng độ CO2 đo được là không đáng kể, trong khi khi bệnh nhân thở ra, nồng độ CO2 cao hơn nhiều do kết quả của các quá trình trao đổi chất. Nói chung, điều này tạo ra một dạng sóng tương tự như sóng vuông đóng vai trò là thước đo vàng tiêu chuẩn của hơi thở.

Có những trường hợp nhất định mà capnography là không điển hình, chẳng hạn như sự giảm biên độ của CO2 cuối kỳ thở ra (có thể chỉ ra suy tim hoặc thuyên tắc phổi [11]). Một ví dụ phổ biến khác là nơi mô hình sóng vuông điển hình trở thành "vây cá mập", và điều này chỉ ra tắc nghẽn đường thở nghiêm trọng như trong trường hợp COPD. Đối với mục đích của công trình này, chúng tôi xem xét có bao nhiêu dạng sóng hô hấp capnography điển hình được mã hóa trong quang thể tích đồ, và không phải các kịch bản y tế khẩn cấp với những thay đổi trong biên độ capnogram.

### **1.3. Autoencoders cho tín hiệu sinh học**

Thông thường, autoencoders được sử dụng như một kỹ thuật học máy không giám sát để học một mã hóa hiệu quả của đầu vào. Autoencoder hoạt động bằng cách nén đầu vào thành một không gian tiềm ẩn, nên đại diện cho thông tin cơ bản đối với tín hiệu đầu vào. Không gian tiềm ẩn này sau đó được lấy mẫu lên trở lại thành đầu ra, mà trong các ứng dụng điển hình nên giống như đầu vào. Các mã hóa hiệu quả của đầu vào tồn tại trong không gian tiềm ẩn sau đó có thể được sử dụng trực tiếp cho phân loại [12], hoặc chính đầu ra có thể được sử dụng cho phân loại [13]. Một phần mở rộng điển hình của nguyên tắc này là khử nhiễu, theo đó đầu vào bị hỏng bởi nhiễu và autoencoder được giao nhiệm vụ khớp đầu ra với phiên bản không bị hỏng của đầu vào. Ví dụ bao gồm khử nhiễu của hình ảnh y tế [14] và khử nhiễu của tín hiệu sinh học như điện tâm đồ (ECG) [15].

Về cơ bản, cấu trúc autoencoder hoạt động bằng cách học một mối quan hệ chính được nén giữa đầu vào và đầu ra. Khi nói đến PPG, rõ ràng có thông tin hô hấp được mã hóa trong tín hiệu. Do đó, chúng tôi giả thuyết rằng bằng cách sử dụng cấu trúc bộ mã hóa-giải mã với quang thể tích đồ làm đầu vào, và huấn luyện đầu ra với một dạng sóng hô hấp tiêu chuẩn vàng dưới dạng capnography, có khả năng chúng ta có thể cô đọng và trích xuất hầu hết thông tin hô hấp liên quan có sẵn trong PPG. Vì việc tái sử dụng này của autoencoder thông thường hiện nhằm trích xuất thông tin tương quan giữa tín hiệu đầu vào và tham chiếu, nên chúng tôi gọi mô hình này là corr-encoder.

## **2. Phương pháp**

### **2.1. Bộ dữ liệu**

Đối với việc kiểm chứng chéo bỏ-một-đối-tượng (leave-one-subject-out cross-validation) ban đầu của mô hình chúng tôi, chúng tôi sử dụng bộ dữ liệu "chuẩn hô hấp" Capnobase [16]. Dữ liệu này bao gồm 42 đối tượng (29 trẻ em và 13 người lớn) được ghi lại trong 8 phút mỗi người trong quá trình thở tự phát hoặc có kiểm soát. Dữ liệu chứa quang thể tích đồ dựa trên ngón tay và capnography được ghi đồng thời, trong đó capnography đóng vai trò là tham chiếu hô hấp tiêu chuẩn vàng. Capnobase bao gồm một phạm vi rộng các tốc độ hô hấp, từ dưới 10 nhịp thở mỗi phút đến trên 40 nhịp thở mỗi phút. Sự phong phú bất thường của các tần số hô hấp cao hơn có thể là do phần lớn bệnh nhân nhi, vì trẻ em có xu hướng có tốc độ hô hấp khi nghỉ ngơi cao hơn.

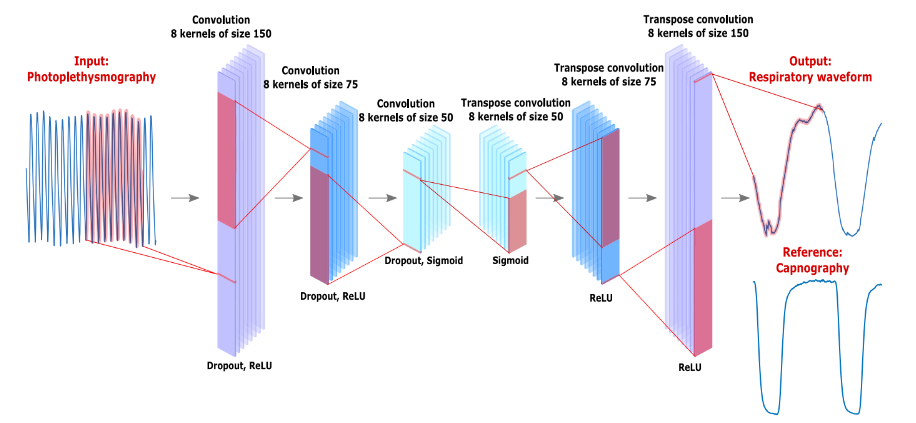
Quang thể tích đồ có thể mất thông tin hô hấp ở tốc độ hô hấp cao hơn do hàm truyền thông thấp giữa hơi thở và PPG [17]. Do đó, việc có nhiều dữ liệu hô hấp tần số cao để kiểm tra, như trong trường hợp bộ dữ liệu Capnobase, là rất quan trọng để đảm bảo tính mạnh mẽ của thuật toán. Capnobase cũng chứa nhiều ví dụ về nhiễu trong đó tín hiệu capnography hoặc PPG bị hỏng. Các nghiên cứu thường loại bỏ các đoạn ghi âm Capnobase có chứa nhiễu [18,19], nhưng trong nghiên cứu này chúng tôi chọn giữ lại các bản ghi này trong phân tích của mình để có cái nhìn thực tế hơn về hiệu suất.

Để kiểm chứng thêm về mô hình, chúng tôi thử nghiệm trên bộ dữ liệu PPG và hô hấp BIDMC [20,21], bao gồm các tín hiệu từ bộ dữ liệu MIMIC-II. Hiệu suất của các thuật toán ước tính tốc độ hô hấp trên bộ dữ liệu BIDMC thường kém hơn so với bộ dữ liệu Capnobase, do sự hiện diện mạnh mẽ của các biến đổi áp suất máu động mạch ở khoảng 0,12 Hz khó phân biệt với các biến đổi hô hấp [20]. Hô hấp trong bộ dữ liệu BIDMC được lấy từ hai đầu dẫn điện tâm đồ được tái sử dụng cho phế phong kế trở kháng, theo dõi chuyển động của lồng ngực trong quá trình hô hấp.

Xét rằng điều này khác về cấu trúc và độ trễ so với capnography, chúng tôi không thể so sánh trực tiếp các dạng sóng hô hấp dự đoán của mình được huấn luyện trên capnography với dạng sóng hô hấp BIDMC về sai số tuyệt đối trung bình. Tuy nhiên, chúng tôi có thể so sánh cả hai dạng sóng trực quan và kiểm tra nội dung tần số để ước tính tốc độ hô hấp. Mục tiêu chính của mô hình được trình bày trong công trình này là trích xuất thông tin chung giữa tham chiếu hô hấp và quang thể tích đồ, độc lập với loại tín hiệu tham chiếu, vẫn là thông tin hô hấp.

Lý do đằng sau cả việc thử nghiệm và đào tạo lại cùng một mô hình trên bộ dữ liệu BIDMC là để kiểm tra khả năng tổng quát hóa của mô hình đối với dữ liệu mới sau khi các tham số của nó đã được tinh chỉnh trên Capnobase. Giống như với bộ dữ liệu Capnobase, chúng tôi sử dụng tất cả dữ liệu trong phân tích của mình và không loại trừ nhiễu. Dữ liệu BIDMC được lấy mẫu lại ở tần số 30 Hz, để phù hợp với tốc độ lấy mẫu của dữ liệu Capnobase đã được giảm mẫu.

PPG thô từ cả bộ dữ liệu Capnobase và BIDMC được lọc thông cao để loại bỏ bất kỳ độ lệch DC nào, với tần số cắt được kiểm soát bởi nhà sản xuất máy đo oxy xung [16]. Như một bước tiếp theo, mỗi cửa sổ PPG được điều chỉnh lại về giữa -1 và 1, và mỗi cửa sổ của capnography và phế phong kế trở kháng được điều chỉnh lại về giữa 0 và 1.



***Hình 1. Kiến trúc học sâu được đề xuất***

***Từ trái sang phải: Dạng sóng quang thể tích đồ được truyền qua 3 lớp tích chập một chiều, mỗi lớp có 8 nhân. Không gian tiềm ẩn thu được sau đó được truyền qua 3 lớp tích chập chuyển vị một chiều, các lớp này lấy mẫu lên không gian tiềm ẩn thành một dạng sóng hô hấp đầu ra.***

***Dưới mỗi lớp, hàm kích hoạt (đơn vị tuyến tính được chỉnh lưu - ReLU hoặc sigmoid) được ghi nhãn, cùng với dropout. Dạng sóng hô hấp đầu ra được mô tả bên cạnh dạng sóng hô hấp tham chiếu, được sử dụng để tính toán sai số bình phương trung bình trong quá trình huấn luyện.***

***Các dạng sóng hiển thị trong hình này là từ một ví dụ đối tượng thử nghiệm trong quá trình kiểm chứng chéo bỏ-một-đối-tượng (leave-one-subject-out cross validation).***

Hình ảnh trình bày kiến trúc mạng nơ-ron sâu "corr-encoder" được sử dụng trong nghiên cứu này để trích xuất dạng sóng hô hấp từ tín hiệu quang thể tích đồ (PPG). Dưới đây là mô tả chi tiết về các thành phần và luồng xử lý:

## **Tổng quan kiến trúc**

Kiến trúc này là một biến thể của autoencoder tích chập, nhưng thay vì tái tạo tín hiệu đầu vào, nó được thiết kế để chuyển đổi tín hiệu PPG thành dạng sóng hô hấp tương tự với capnography (tham chiếu tiêu chuẩn vàng).

## **Các thành phần chính**

1. **Đầu vào (Input)**:
   * Tín hiệu Quang thể tích đồ (Photoplethysmography) - được thể hiện bằng đường sóng màu xanh ở bên trái
2. **Phần mã hóa (Encoder)**:
   * Lớp 1: Tích chập (Convolution) với 8 nhân (kernels) kích thước 150, sau đó là Dropout và hàm kích hoạt ReLU
   * Lớp 2: Tích chập với 8 nhân kích thước 75, sau đó là Dropout và hàm kích hoạt ReLU
   * Lớp 3: Tích chập với 8 nhân kích thước 50, sau đó là Dropout và hàm kích hoạt Sigmoid
3. **Phần giải mã (Decoder)**:
   * Lớp 1: Tích chập chuyển vị (Transpose convolution) với 8 nhân kích thước 50, sau đó là hàm kích hoạt Sigmoid
   * Lớp 2: Tích chập chuyển vị với 8 nhân kích thước 75, sau đó là hàm kích hoạt ReLU
   * Lớp 3: Tích chập chuyển vị với 8 nhân kích thước 150, sau đó là hàm kích hoạt ReLU
4. **Đầu ra (Output)**:
   * Dạng sóng hô hấp (Respiratory waveform) được tạo ra từ mô hình (đường màu tím/đỏ)
   * Được so sánh với dạng sóng tham chiếu Capnography (đường màu xanh ở góc dưới bên phải)

## **Nguyên lý hoạt động**

* **Phần mã hóa**: Nén thông tin từ tín hiệu PPG đầu vào thông qua các lớp tích chập liên tiếp, trích xuất các đặc trưng liên quan đến hô hấp
* **Phần giải mã**: Từ đặc trưng đã trích xuất, tái tạo không phải tín hiệu PPG ban đầu mà là dạng sóng hô hấp tương ứng
* **Các kích thước nhân khác nhau**: Cho phép mạng nắm bắt các mẫu ở các quy mô thời gian khác nhau, từ ngắn đến dài

Mục tiêu của kiến trúc này là trích xuất thông tin hô hấp đã được mã hóa trong tín hiệu PPG và chuyển đổi nó thành dạng sóng tương tự với dạng sóng capnography tiêu chuẩn. Đây là lý do tại sao mô hình được gọi là "corr-encoder" - nó tập trung vào việc trích xuất thông tin tương quan (correlated) giữa PPG và tín hiệu hô hấp tham chiếu.

## **2.2. Mô hình học sâu**

Một cấu trúc bộ mã hóa-giải mã tích chập một chiều đã được triển khai trong Pytorch [22]. Tổng quan về cấu trúc mô hình được thể hiện trong Hình 1. Mô hình bao gồm 3 lớp tích chập "mã hóa" với mỗi lớp có 8 nhân kích thước lần lượt là 150, 75 và 50 mẫu, sau đó được phản chiếu với 3 lớp tích chập chuyển vị lấy mẫu lên "giải mã". Cả đầu vào PPG và ước tính dạng sóng hô hấp đầu ra đều có độ dài 288 mẫu, tương ứng với 9,6 giây dữ liệu. Thời lượng 9,6 giây đơn giản là kết quả của việc chia tổng cộng 480 giây dữ liệu cho mỗi đối tượng thành 50 phân đoạn bằng nhau. Mặc dù điều này hơi tùy ý, nhưng điều quan trọng là phải có độ dài cửa sổ chứa ít nhất một chu kỳ hô hấp hoàn chỉnh.

Hàm kích hoạt tuyến tính được chỉnh lưu (ReLU) được áp dụng cho hai lớp đầu tiên, với hàm kích hoạt sigmoid được áp dụng cho các lớp trong cùng. Một dropout 0,5 được áp dụng cho các lớp mã hóa. Đầu vào và đầu ra của lớp đầu tiên đều được đệm (padding) thêm 20, và đầu ra của lớp thứ hai được đệm thêm 10. Điều này lại được phản chiếu trong các lớp tích chập chuyển vị trong bộ giải mã. Mô hình được huấn luyện bằng cách tối thiểu hóa sai số bình phương trung bình giữa đầu ra và capnography tiêu chuẩn vàng. Tối ưu hóa Adam được triển khai với tốc độ học là 1 × 10^-3.

Các tham số của mô hình được tối ưu hóa độc quyền trên bộ dữ liệu Capnobase, với việc huấn luyện diễn ra trên 39 đối tượng đầu tiên và kiểm tra trên 3 đối tượng cuối cùng có phạm vi chất lượng tín hiệu và tốc độ thở tốt. Khi chọn số lượng lớp, có một cải thiện hiệu suất nhỏ khi tăng từ 2 lớp mã hóa và giải mã lên 3 lớp mã hóa và giải mã, nhưng 4 lớp không dẫn đến tăng hiệu suất và có xu hướng quá khớp (over fit).

Khi chọn số lượng nhân cho mỗi lớp, có thể thấy trong Hình 2(c) rằng không có sự gia tăng đáng kể về sai số tuyệt đối trung bình kiểm tra trung vị khi sử dụng hơn 8 nhân cho mỗi lớp, và có sự gia tăng trong khoảng tứ phân vị cho thấy độ ổn định của mô hình giảm sau điểm này. Khi chọn kích thước nhân, hiệu suất phần lớn không bị ảnh hưởng khi thay đổi kích thước nhân lớp 1 giữa 100 và 150, nhưng kích thước nhân lớp 1 vượt quá 150 dẫn đến sự suy giảm hiệu suất đáng kể.

Chúng tôi đã chọn duy trì kích thước nhân lớn hơn trong lớp đầu vào so với thông thường để cung cấp khả năng giải thích khi so sánh với kích thước của tín hiệu đầu vào và đầu ra. Quan trọng là, hiệu suất được cải thiện đáng kể nhờ có kích thước nhân giảm dần theo các lớp tiếp theo. Phần đệm (padding) được giữ ở mức tối thiểu để cải thiện thời gian tính toán nhưng vẫn khác không, vì không có đệm nào dẫn đến hiệu suất giảm.

Dropout cao trong các lớp mã hóa đã cải thiện hiệu suất trên dữ liệu kiểm tra, vì nó cải thiện khả năng khái quát hóa của mô hình đối với các đầu vào khác nhau, nhưng dropout có ảnh hưởng bất lợi đáng kể đến độ chính xác khi áp dụng cho các lớp giải mã. Tốc độ học tăng lên 1 × 10^-2 gặp khó khăn trong việc hội tụ, và tốc độ học giảm xuống 1 × 10^-4 mất nhiều thời gian hơn để hội tụ đến một giải pháp, và do đó tốc độ học 1 × 10^-3 được xác định là điểm cân bằng tốt nhất.

Dừng sớm (early stopping) cũng được triển khai bằng cách chấm dứt huấn luyện tại 80 epoch để tránh quá khớp. Mặc dù không đặc biệt cần thiết, ReLU được triển khai trong các lớp nông xét rằng tín hiệu đầu ra được chuẩn hóa giữa 0 và 1, và dẫn đến sự giảm nhỏ trong sai số bình phương trung bình. Việc bổ sung các hàm kích hoạt sigmoid trong các lớp bên trong đã cải thiện đáng kể về độ ổn định vì chúng ngăn các giá trị tăng vọt, do đặc tính đầu ra có giới hạn của chúng.

## **2.3. Đánh giá mô hình**

Mô hình đầu tiên được kiểm tra trên bộ dữ liệu Capnobase và phương pháp kiểm chứng chéo bỏ-một-đối-tượng (leave-one-subject-out cross-validation) đã được triển khai. Khi kiểm tra trên các đối tượng chưa từng thấy, thay vì dữ liệu quang thể tích đồ của đối tượng được chia thành 50 phân đoạn riêng biệt, dữ liệu được phân đoạn với một cửa sổ trượt 9,6 giây (288 mẫu) được dịch chuyển 1 giây mỗi lần. Điều này tạo ra 471 phân đoạn kiểm tra cho mỗi đối tượng. Các phân đoạn đầu ra của mô hình sau đó được kết hợp và lấy trung bình, trong một kích thước cửa sổ đánh giá nhất định, để đưa ra ước tính dạng sóng hô hấp mượt mà hơn.

Lý do đằng sau việc kiểm tra với các cửa sổ chồng lấp là vì với độ phức tạp thấp của mô hình, một cửa sổ 9,6 giây của dạng sóng hô hấp có thể được tính toán trong một phần nhỏ của mili giây. Việc lấy trung bình các cửa sổ liên tiếp cải thiện độ mạnh mẽ của dạng sóng đầu ra với gần như không có chi phí thêm. Hơn nữa, trong một triển khai thế giới thực, 9,6 giây thông tin hô hấp trước đó sẽ được tính toán với tốc độ làm mới nhanh theo cách tương tự, thay vì đợi thêm 9,6 giây nữa để chuyển sang mỗi cửa sổ mới. Vì những lý do này, chiến lược này đã được triển khai khi kiểm tra mô hình. Khi huấn luyện mô hình, cùng một chiến lược tạo đầu vào chồng lấp là không cần thiết do khối lượng lớn dữ liệu đã có sẵn, và sẽ làm tăng thời gian huấn luyện.

Các dạng sóng hô hấp đầu ra đầu tiên được đánh giá bằng cách so sánh sai số tuyệt đối trung bình với dạng sóng capnography tham chiếu. Kết quả sai số tuyệt đối trung bình cũng được so sánh với phương pháp hồi quy Bình phương Tối thiểu Riêng phần (Partial Least Squares - PLS) [23], một cách tiêu chuẩn để tìm phép chiếu lên không gian tiềm ẩn, được huấn luyện và kiểm tra theo cùng một cách chính xác bằng MATLAB. Trong trường hợp này, PLS được triển khai với 26 bậc tự do, được phát hiện là mang lại sai số dạng sóng tuyệt đối trung bình trung vị tối ưu trên dữ liệu kiểm tra.

Phương pháp Bình phương Tối thiểu Riêng phần nhằm dự đoán một đầu ra từ một đầu vào nhất định bằng cách biến đổi đầu vào thông qua phép nhân với các vectơ tải đầu vào, và biến đổi đầu ra với các vectơ tải đầu ra. Thuật toán được thiết kế sao cho một đầu vào được biến đổi có hiệp phương sai tối đa với đầu ra được biến đổi tương ứng. Theo nghĩa này, có thể vẽ ra sự tương đồng giữa các nhân tích chập của mô hình corr-encoder và các vectơ tải của PLS. So sánh này rất quan trọng vì chúng tôi đang sử dụng cấu trúc bộ mã hóa-giải mã theo cách tương tự như những gì PLS được thiết kế để thực hiện. Phương pháp Bình phương Tối thiểu Riêng phần được đánh giá theo cùng một cách với mô hình học sâu, thông qua huấn luyện bỏ một đối tượng và kiểm tra, và lấy trung bình các đầu ra kiểm tra chồng lấp.

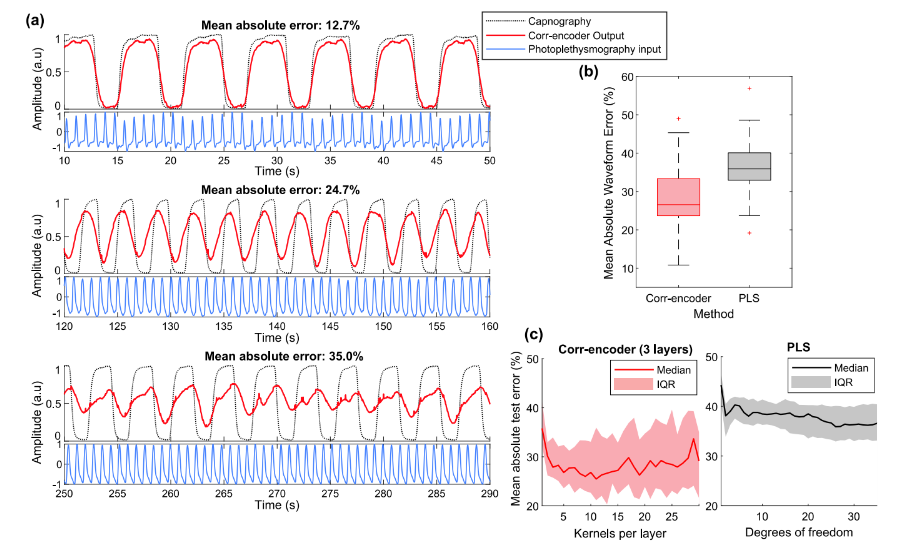
Một phép quét tham số cho sai số dạng sóng tuyệt đối trung bình so với bậc tự do được cung cấp cho PLS trong Hình 2(c), bên cạnh phân tích tương ứng cho số lượng nhân trên mỗi lớp cho corr-encoder. Điều thú vị cần lưu ý là tổng số nhân tối ưu cho corr-encoder từ 8 đến 11 nhân cho mỗi lớp (dẫn đến tổng cộng 24 đến 33 nhân mã hóa) phù hợp với bậc tự do tối ưu là 26 cho Partial Least Squares.

Đối với bộ dữ liệu Capnobase, các dạng sóng đầu ra được đánh giá cho cả ước tính tốc độ hô hấp và chu kỳ hoạt động (tỷ lệ giữa thời gian hít vào và thời gian thở ra). Để ước tính tốc độ hô hấp, phép biến đổi Fourier nhanh (FFT) được tính toán và tần số tương ứng với công suất tối đa được trích xuất. Tần số này tính bằng Hertz sau đó được nhân với 60 để cho tốc độ hô hấp tính bằng nhịp thở mỗi phút. Ước tính tốc độ hô hấp được so sánh với tốc độ hô hấp tiêu chuẩn vàng, được tính toán thủ công trong bộ dữ liệu từ dạng sóng capnography.

FFT được chọn để đánh giá mô hình của chúng tôi thay vì đếm thủ công để tính toán tốc độ hô hấp được tự động, làm cho kết quả thực tế hơn cho việc triển khai trong thế giới thực. Hơn nữa, tính toán thủ công được chọn thay vì FFT để so sánh với tín hiệu tham chiếu để tốc độ hô hấp tham chiếu vẫn có thể được coi là chân lý nền. Để cho phép so sánh với các mô hình khác nhau trong tài liệu, cả sai số tốc độ hô hấp tuyệt đối trung vị trên tất cả các cửa sổ và sai số tuyệt đối trung bình trung vị trên tất cả các đối tượng đều được tính toán.

Đối với ước tính chu kỳ hoạt động, xét rằng các dạng sóng được huấn luyện đến biên độ giữa 0 và 1, chu kỳ hoạt động hít vào được ước tính là số lượng mẫu dưới 0,5 được chia cho tổng số mẫu. Sau đó, con số này được nhân với 100 để cho ra tỷ lệ phần trăm. Cùng một thao tác được thực hiện trên capnography tiêu chuẩn vàng để so sánh. Cả sai số tuyệt đối trung bình và trung vị trong chu kỳ hoạt động đều được trình bày.

Đối với bộ dữ liệu BIDMC, chỉ ước tính tốc độ hô hấp được sử dụng làm thước đo đánh giá, vì dạng sóng hô hấp tiêu chuẩn vàng là phế phong kế trở kháng chứ không còn là capnography. Hai mô hình huấn luyện đã được triển khai để đánh giá. Thứ nhất, mô hình được huấn luyện trên toàn bộ bộ dữ liệu Capnobase và sau đó được kiểm tra trên bộ dữ liệu BIDMC. Thứ hai, mô hình được huấn luyện trên bộ dữ liệu Capnobase, và sau đó được huấn luyện lại trên bộ dữ liệu BIDMC theo cách bỏ một đối tượng. Trong tất cả các trường hợp trong công trình này, mô hình được đánh giá trên các đối tượng kiểm tra chưa từng thấy.



***Hình 2. Các dạng sóng hô hấp đầu ra ví dụ và sai số tuyệt đối trung bình từ mô hình đề xuất.***

***(a) Các dạng sóng đầu ra (màu đỏ - nét liền), dạng sóng capnography tiêu chuẩn vàng (màu đen - đứt nét) và dạng sóng quang thể tích đồ đầu vào (màu xanh - nét liền) cho ba kịch bản khác nhau (từ trên xuống dưới) gồm tái tạo gần như hoàn hảo, tái tạo điển hình và tái tạo kém.***

***(b) Biểu đồ hộp về sai số dạng sóng tuyệt đối trung bình khi kiểm tra trên tất cả 42 đối tượng của bộ dữ liệu Capnobase, cho khung học sâu đề xuất (màu đỏ) và hồi quy Bình phương Tối thiểu Riêng phần (màu xám).***

***(c) Kết quả của các phép quét tham số cho cả khung corr-encoder và PLS, thể hiện trung vị và khoảng tứ phân vị của sai số dạng sóng tuyệt đối trung bình trên các đối tượng kiểm tra, theo số lượng nhân trên mỗi lớp đối với corr-encoder và bậc tự do đối với PLS.***

# **Giải thích chi tiết Hình 2**

## **Phần (a) - So sánh dạng sóng hô hấp dự đoán**

Phần (a) trình bày ba ví dụ về dạng sóng hô hấp ở các mức độ chính xác khác nhau:

1. **Panel trên cùng (Sai số 12.7%)**:
   * Đường đen đứt (Capnography): Tín hiệu hô hấp tham chiếu tiêu chuẩn vàng
   * Đường đỏ (Corr-encoder): Dạng sóng hô hấp được dự đoán từ mô hình
   * Đường xanh dưới: Tín hiệu Photoplethysmography đầu vào
   * Tại mức sai số thấp nhất (12.7%), dạng sóng dự đoán gần như trùng khớp hoàn toàn với tham chiếu
2. **Panel giữa (Sai số 24.7%)**:
   * Hiển thị một trường hợp có chất lượng dự đoán trung bình
   * Dạng sóng dự đoán vẫn bám sát tần số hô hấp, nhưng có sự khác biệt về biên độ và hình dạng
   * Tín hiệu PPG đầu vào có tần số cao hơn nhiều so với sóng hô hấp
3. **Panel dưới cùng (Sai số 35.0%)**:
   * Thể hiện trường hợp dự đoán kém nhất
   * Dạng sóng dự đoán vẫn nắm bắt được tần số hô hấp cơ bản, nhưng biên độ và hình dạng khác biệt đáng kể
   * Đỉnh và đáy của sóng dự đoán không khớp hoàn toàn với tham chiếu

## **Phần (b) - So sánh hiệu suất giữa các phương pháp**

Phần (b) trình bày biểu đồ hộp so sánh sai số tuyệt đối trung bình của dạng sóng giữa hai phương pháp:

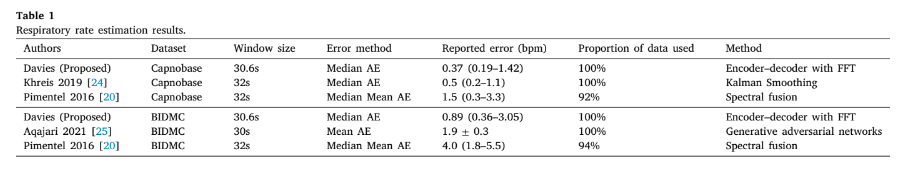
* **Corr-encoder (bên trái)**:
  + Giá trị trung vị thấp hơn (khoảng 28%)
  + Phạm vi tứ phân vị (IQR) hẹp hơn (từ khoảng 23% đến 33%)
  + Sai số thấp nhất khoảng 12%, cao nhất khoảng 55%
* **PLS (Partial Least Squares) (bên phải)**:
  + Giá trị trung vị cao hơn (khoảng 37%)
  + Phạm vi tứ phân vị rộng hơn (từ khoảng 33% đến 42%)
  + Hiệu suất kém hơn và ít ổn định hơn so với Corr-encoder

## **Phần (c) - Phân tích tham số của mô hình**

Phần (c) trình bày sự ảnh hưởng của các tham số đến hiệu suất của hai phương pháp:

1. **Corr-encoder (bên trái)**:
   * Đồ thị cho thấy sai số tuyệt đối trung bình thay đổi theo số lượng nhân trên mỗi lớp
   * Đường đỏ thể hiện giá trị trung vị, vùng hồng thể hiện khoảng tứ phân vị (IQR)
   * Hiệu suất tối ưu đạt được ở khoảng 8-11 nhân trên mỗi lớp
   * Tăng số lượng nhân trên 15 không cải thiện hiệu suất, thậm chí làm tăng độ biến thiên (IQR rộng hơn)
2. **PLS (bên phải)**:
   * Đồ thị cho thấy sai số tuyệt đối trung bình thay đổi theo số bậc tự do
   * Đường đen thể hiện giá trị trung vị, vùng xám thể hiện khoảng tứ phân vị
   * Hiệu suất tối ưu đạt được ở khoảng 26 bậc tự do
   * Đáng chú ý là con số này gần với tổng số nhân tối ưu của Corr-encoder (8-11 nhân × 3 lớp = 24-33 nhân)

Kết luận: Hình này chứng minh hiệu quả vượt trội của mô hình Corr-encoder so với phương pháp PLS truyền thống trong việc trích xuất dạng sóng hô hấp từ tín hiệu PPG. Mô hình có thể tái tạo dạng sóng hô hấp từ tín hiệu PPG với độ chính xác cao và ổn định hơn, đặc biệt khi số lượng nhân được chọn phù hợp.



# **Bảng 1: Giải thích chi tiết kết quả ước tính tốc độ hô hấp**

Bảng 1 trình bày kết quả so sánh hiệu suất của các phương pháp khác nhau trong việc ước tính tốc độ hô hấp trên hai bộ dữ liệu: Capnobase và BIDMC.

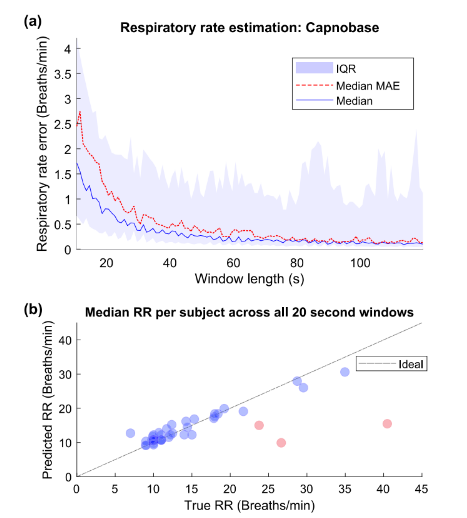
## **Cấu trúc bảng:**

1. **Cột Authors (Tác giả)**:
   * Liệt kê các nhóm nghiên cứu, trong đó "Davies (Proposed)" là mô hình corr-encoder được đề xuất trong nghiên cứu này
   * Khreis 2019 [24], Pimentel 2016 [20] và Asgari 2021 [25] là các nghiên cứu trước đó để so sánh
2. **Cột Dataset (Bộ dữ liệu)**:
   * Capnobase: bộ dữ liệu sử dụng capnography làm tham chiếu chuẩn
   * BIDMC: bộ dữ liệu thứ hai với tham chiếu hô hấp từ phế phong kế trở kháng
3. **Cột Window size (Kích thước cửa sổ)**:
   * Thời gian (giây) của mỗi đoạn dữ liệu được xử lý (30-32 giây)
4. **Cột Error method (Phương pháp đo lỗi)**:
   * Median AE: Sai số tuyệt đối trung vị
   * Mean AE: Sai số tuyệt đối trung bình
   * Median Mean AE: Trung vị của sai số tuyệt đối trung bình giữa các đối tượng
5. **Cột Reported error (bpm) (Lỗi được báo cáo - nhịp thở/phút)**:
   * Giá trị sai số theo đơn vị nhịp thở mỗi phút
   * Số trong ngoặc biểu thị khoảng giá trị hoặc độ lệch chuẩn
6. **Cột Proportion of data used (Tỷ lệ dữ liệu được sử dụng)**:
   * Phần trăm dữ liệu được sử dụng trong phân tích (một số phương pháp loại bỏ dữ liệu có nhiễu)
7. **Cột Method (Phương pháp)**:
   * Mô tả kỹ thuật được sử dụng trong mỗi nghiên cứu

## **Những phát hiện chính:**

1. **Hiệu suất trên bộ dữ liệu Capnobase**:
   * Mô hình encoder-decoder đề xuất đạt sai số thấp nhất (0.37 bpm)
   * Phương pháp Kalman Smoothing của Khreis có sai số 0.5 bpm
   * Phương pháp Spectral fusion của Pimentel có sai số cao nhất (1.5 bpm)
2. **Hiệu suất trên bộ dữ liệu BIDMC**:
   * Mô hình encoder-decoder vẫn đạt sai số thấp nhất (0.89 bpm)
   * Phương pháp Generative adversarial networks của Asgari có sai số 1.9 bpm
   * Phương pháp Spectral fusion của Pimentel có sai số rất cao (4.0 bpm)
3. **So sánh giữa các bộ dữ liệu**:
   * Tất cả các phương pháp đều hoạt động tốt hơn trên Capnobase so với BIDMC
   * Điều này phù hợp với phần mô tả trước đó về BIDMC là bộ dữ liệu khó hơn do có nhiễu trong các tín hiệu
4. **Tỷ lệ dữ liệu sử dụng**:
   * Phương pháp đề xuất và của Khreis sử dụng 100% dữ liệu
   * Các phương pháp của Pimentel chỉ sử dụng 92-94% dữ liệu (có thể đã loại bỏ dữ liệu nhiễu)

Bảng này chứng minh hiệu quả vượt trội của mô hình corr-encoder so với các phương pháp hiện có trong việc ước tính chính xác tốc độ hô hấp từ tín hiệu PPG, đặc biệt khi xử lý dữ liệu có nhiễu như trong bộ dữ liệu BIDMC.



# **Hình 3. Kết quả ước tính tốc độ hô hấp cho bộ dữ liệu Capnobase.**

(a) Sai số ước tính tốc độ hô hấp trên tất cả các độ dài cửa sổ, với giá trị trung vị trên tất cả các cửa sổ được thể hiện bằng màu xanh, và sai số tuyệt đối trung bình trung vị trên tất cả các đối tượng được thể hiện bằng màu đỏ.

(b) Biểu đồ phân tán của tốc độ hô hấp trung vị dự đoán so với tốc độ hô hấp trung vị thực tế cho mỗi đối tượng, được tính toán sử dụng tất cả các cửa sổ ước tính 20 giây cho một đối tượng cụ thể. Ba trong số 42 đối tượng mà mô hình hoạt động kém được đánh dấu màu đỏ.

# **Phân tích chi tiết Hình 3: Kết quả ước tính tốc độ hô hấp**

## **Phần (a) - Ảnh hưởng của kích thước cửa sổ đánh giá**

Biểu đồ phần (a) thể hiện sự thay đổi trong độ chính xác ước tính tốc độ hô hấp trên bộ dữ liệu Capnobase khi thay đổi kích thước cửa sổ đánh giá.

**Các thành phần chính:**

* **Đường xanh dương liền**: Thể hiện giá trị trung vị của sai số
* **Đường đỏ đứt**: Thể hiện sai số tuyệt đối trung bình trung vị (Median MAE)
* **Vùng màu tím nhạt**: Biểu thị khoảng tứ phân vị (IQR), thể hiện độ phân tán của sai số

**Xu hướng quan trọng:**

* Sai số giảm mạnh khi kích thước cửa sổ tăng từ 10s đến khoảng 60s
* Với cửa sổ dài 20s, sai số khoảng 1-1.5 nhịp thở/phút
* Khi cửa sổ đạt 60s trở lên, sai số giảm xuống dưới 0.5 nhịp thở/phút và ổn định
* Khoảng tứ phân vị cũng thu hẹp theo kích thước cửa sổ, cho thấy độ ổn định cao hơn với cửa sổ dài

Biểu đồ này cho thấy một sự đánh đổi giữa thời gian phân tích (kích thước cửa sổ) và độ chính xác: cửa sổ càng dài, ước tính càng chính xác và ổn định.

## **Phần (b) - So sánh tốc độ hô hấp dự đoán với giá trị thực**

Biểu đồ phần (b) là biểu đồ phân tán thể hiện mối quan hệ giữa tốc độ hô hấp dự đoán từ mô hình (trục y) và tốc độ hô hấp thực tế (trục x) đo trên cửa sổ 20 giây.

**Đặc điểm quan trọng:**

* **Đường chéo đứt**: Biểu thị dự đoán lý tưởng (khi giá trị dự đoán = giá trị thực)
* **Điểm màu xanh lam**: Đại diện cho phần lớn các đối tượng, nằm rất gần đường lý tưởng
* **Điểm màu đỏ**: Đại diện cho các trường hợp dự đoán sai (có thể là các đối tượng với dữ liệu nhiễu)

**Kết quả chính:**

* Phần lớn các điểm nằm gần đường lý tưởng, đặc biệt trong khoảng tốc độ hô hấp phổ biến (5-20 nhịp thở/phút)
* Mô hình hoạt động rất tốt với tốc độ hô hấp thông thường
* Có một vài trường hợp dự đoán sai (điểm đỏ) đặc biệt ở các tốc độ hô hấp cao (>25 nhịp thở/phút)
* Phạm vi tốc độ hô hấp trong dữ liệu rất rộng, từ khoảng 5 đến 45 nhịp thở/phút

Biểu đồ phần (b) xác nhận rằng mô hình corr-encoder có khả năng dự đoán chính xác tốc độ hô hấp cho phần lớn đối tượng, nhưng có thể gặp khó khăn với một số trường hợp đặc biệt, đặc biệt là khi tốc độ hô hấp cao bất thường.

## **3. Kết quả**

### **3.1. Ước tính dạng sóng**

Các dạng sóng đầu ra kiểm tra điển hình được vẽ đồ thị so với capnography tiêu chuẩn vàng cho ba ví dụ về trường hợp tốt (MAE 13%), điển hình (MAE 25%) và kém (MAE 35%) trong Hình 2(a). Có thể thấy rằng ngay cả khi sự khác biệt giữa các dạng sóng là lớn trong ví dụ với MAE 35%, các thành phần tần số cơ bản vẫn được nắm bắt.

Sai số tuyệt đối trung bình trung vị của các đầu ra mô hình trên các đối tượng kiểm tra chưa từng thấy là 27%, với khoảng tứ phân vị từ 24% đến 33%. Con số này so với MAE trung vị 36% và khoảng tứ phân vị từ 33% đến 40% đối với phương pháp hồi quy Bình phương Tối thiểu Riêng phần (Partial Least Squares), như được thể hiện trong Hình 2(b).

# **3.2. Ước tính tần số hô hấp**

Để so sánh với các nghiên cứu khác về ước tính tần số hô hấp (RR) trong tài liệu, các sai số ước tính RR được trình bày theo cả sai số tuyệt đối trung vị (mAE) trên tất cả các cửa sổ thời gian, và sai số tuyệt đối trung bình trung vị (mMAE) trên tất cả các đối tượng nghiên cứu. Đối với cửa sổ thời gian 30,6 giây, mAE đạt 0,37 nhịp thở mỗi phút, và mMAE đạt 0,50 nhịp thở mỗi phút. Đối với cửa sổ thời gian 60,6 giây, kết quả được cải thiện nhẹ với mAE là 0,16 nhịp thở mỗi phút và mMAE là 0,40 nhịp thở mỗi phút.

Các giá trị MAE và mMAE được biểu diễn cho tất cả các cửa sổ thời gian từ 10 giây đến 120 giây trong Hình 3(a). Trong Hình 3(b), sai số trung bình trung vị được biểu diễn cho từng đối tượng với cửa sổ ước tính 20,6 giây, với đường đứt nét biểu thị dự đoán lý tưởng để so sánh. Cửa sổ thời gian ngắn hơn này được chọn để minh họa rằng ngay cả với độ dài cửa sổ ngắn hơn nhiều so với những gì đã được trình bày trong tài liệu trước đây, việc ước tính tần số hô hấp vẫn chính xác trên toàn bộ phạm vi tần số hô hấp và không có độ lệch. Ngoài ra, kết quả cũng chỉ ra rằng việc ước tính kém chính xác ở 3 trong số 42 đối tượng (được đánh dấu màu đỏ), và sai số này vẫn tồn tại bất kể kích thước cửa sổ thời gian nào được áp dụng.

# **3.3. Thời gian hô hấp**

Khi so sánh thời gian ước tính dành cho hít vào, được tính theo tỷ lệ của toàn bộ chu kỳ hô hấp, giữa kết quả từ mô hình học sâu và tiêu chuẩn vàng capnography, hệ số tương quan Pearson trung vị đạt 0,16 (P. = 1,8 × 10^-3). Sự tương quan dương yếu này cho thấy phần lớn các ước tính dạng sóng hô hấp từ photoplethysmography ngón tay không thể nắm bắt đầy đủ tỷ lệ thời gian giữa hít vào và thở ra.

# **3.4. Kiểm chứng trên BIDMC**

Khi huấn luyện mô hình trên bộ dữ liệu Capnobase và kiểm tra trên bộ dữ liệu BIDMC chưa từng thấy, việc ước tính tần số hô hấp nhìn chung chính xác với MAE và mMAE lần lượt là 1,74 và 3,02 nhịp thở mỗi phút với cửa sổ 30,6 giây, và MAE và mMAE lần lượt là 1,08 và 2,06 nhịp thở mỗi phút, sử dụng cửa sổ 60,6 giây. Khi huấn luyện mô hình trên bộ dữ liệu Capnobase, và sau đó huấn luyện lại trên bộ dữ liệu BIDMC, các kết quả được cải thiện với MAE và mMAE lần lượt là 0,89, 1,50 và 0,58, 1,01 nhịp thở mỗi phút với cửa sổ 30,6 giây và 60,6 giây. Sai số ước tính tần số hô hấp được hiển thị cho tất cả các cửa sổ từ 10 giây đến 120 giây trong Hình 4(a).

Hơn nữa, Hình 4(b) cho thấy với một ví dụ dạng sóng điển hình, ngay cả khi không huấn luyện trên BIDMC, phần lớn thông tin hô hấp vẫn được nắm bắt. Việc huấn luyện lại trên BIDMC được chứng minh có tác dụng thay đổi hình thái dạng sóng đầu ra để giống với impedance pneumography, được sử dụng làm tham chiếu trong trường hợp này, thay vì capnography. Để mục đích so sánh, mô hình corr-encoder cũng được huấn luyện độc quyền trên BIDMC, dẫn đến MAE và mMAE lần lượt là 0,90, 1,58 và 0,81, 1,36 nhịp thở mỗi phút với cửa sổ 30,6 giây và 60,6 giây. Những kết quả này kém hơn so với việc mô hình được huấn luyện trước trên Capnobase và huấn luyện lại trên BIDMC, cho thấy bộ dữ liệu Capnobase chứa thông tin hô hấp bổ sung không được nắm bắt bởi BIDMC, và do đó việc huấn luyện trên cả hai bộ dữ liệu giúp mô hình tổng quát hóa tốt hơn đối với các đối tượng chưa từng thấy của BIDMC.

# **4. Thảo luận**

Đối với việc tái tạo dạng sóng capnography, cấu trúc encoder-decoder được trình bày ở đây vượt trội hơn so với phương pháp hồi quy Partial Least Squares tiêu chuẩn. Nghiên cứu chỉ ra rằng mô hình thành công trong việc tái tạo gần như hoàn hảo về biên độ, pha và thời gian hô hấp chỉ ở một tỷ lệ nhỏ các đối tượng. Trong đa số trường hợp, thông tin biên độ không được nắm bắt hoàn toàn, dẫn đến sai số tuyệt đối trung bình lớn hơn 0,23 ở ba phần tư các đối tượng thử nghiệm. Tuy nhiên, có thể thấy rằng trong các dạng sóng mà thông tin biên độ không được nắm bắt, dẫn đến sai số dạng sóng tuyệt đối trung bình cao hơn, thông tin tần số chủ đạo của dạng sóng hô hấp vẫn được nắm bắt một cách hiệu quả.

Điều này được phản ánh thêm trong khả năng của mô hình đối với việc ước tính chính xác tần số hô hấp, thu được đơn giản bằng cách lấy biến đổi Fourier nhanh của dạng sóng đầu ra. Bảng 1 cho thấy khi so sánh với các thuật toán ước tính tần số hô hấp khác, dạng sóng được tạo ra bởi mô hình của chúng tôi có độ chính xác tương tự như các phương pháp tiên tiến nhất trong cả bộ dữ liệu Capnobase và BIDMC. Khi thực hiện các so sánh này, chúng tôi đã cẩn thận đưa vào các ví dụ gần đây sử dụng tất cả dữ liệu trong một bộ dữ liệu nhất định, cũng như một thuật toán đã được áp dụng trên cả hai bộ dữ liệu. So sánh với các thuật toán được triển khai trên bộ dữ liệu BIDMC đặc biệt quan trọng, xét đến độ khó mà nó mang lại cho việc ước tính tần số hô hấp thông qua sự hiện diện của các biến đổi tần số thấp không liên quan đến hô hấp trong PPG [20].

Mặc dù sai số của chúng tôi thấp hơn Pimentel et al. [20] trên bộ dữ liệu Capnobase, điều quan trọng cần lưu ý là Khreis et al., những người sử dụng phương pháp dựa trên làm mịn Kalman [24], đã đạt được độ chính xác tương đương với phương pháp của chúng tôi trên bộ dữ liệu Capnobase. Thuật toán làm mịn Kalman dựa trên ma trận chuyển đổi trạng thái được hình thành từ 32 giây dữ liệu trước đó, có tác dụng làm mịn các biến động nhanh trong tốc độ hô hấp. Mặc dù hạn chế này không ảnh hưởng đến độ chính xác trên bộ dữ liệu Capnobase, nơi tốc độ hô hấp thường ổn định trong suốt thời gian ghi âm 8 phút, nó sẽ ảnh hưởng đến độ chính xác trong các ứng dụng có sự thay đổi đột ngột về tốc độ hô hấp.

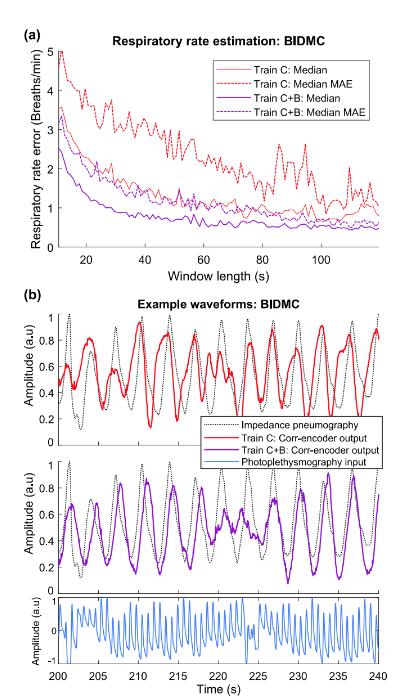
Hạn chế này không áp dụng cho mô hình corr-encoder đề xuất của chúng tôi, vốn có thể phản ứng ngay lập tức với những thay đổi đột ngột về tốc độ hô hấp. Trên bộ dữ liệu BIDMC, Aqajari et al. đã đạt được độ chính xác tương đương với phương pháp của chúng tôi bằng cách sử dụng mạng đối nghịch sản sinh chu kỳ (cycleGAN) [25]. Kiến trúc cycleGAN cụ thể được triển khai bao gồm một mạng sản sinh gồm 16 lớp tích chập và hơn 1600 nhân [26] và một mạng phân biệt gồm 4 lớp và hơn 900 nhân [27], khi kết hợp là một mô hình 20 lớp với hơn 2500 nhân. Điều này tương phản với mô hình đơn giản hơn nhiều của chúng tôi, có 6 lớp với tổng cộng 48 nhân. Với độ phức tạp tương đối thấp của kiến trúc của chúng tôi, có thể huấn luyện mô hình trong chưa đầy 6 phút trên CPU, so với kiến trúc cycleGAN khi áp dụng cho hình ảnh có thể mất hàng giờ để huấn luyện trên GPU [26].

Phân tích dạng sóng hô hấp vượt ra ngoài tần số hô hấp, chẳng hạn như xem xét tỷ lệ thời gian giữa hít vào và thở ra (chu kỳ nhiệm vụ), có thể quan trọng để phân biệt các bệnh hô hấp [1]. Ước tính thành công chu kỳ nhiệm vụ hít vào đã được thực hiện trước đây từ PPG trong tai nhằm mục đích phân loại bệnh phổi tắc nghẽn mãn tính [8]. Ở đây chúng tôi chỉ ra rằng với sự kết hợp của phương pháp được trình bày trong công trình này và bộ dữ liệu Capnobase, việc trích xuất chu kỳ nhiệm vụ chính xác phần lớn không thành công, chỉ đạt được mối tương quan dương yếu giữa chu kỳ nhiệm vụ thực tế và ước tính của chúng tôi. Hàm truyền bộ lọc thông thấp giữa hô hấp và PPG [17] thường dẫn đến mất chi tiết chu kỳ nhiệm vụ [8]. Các biến đổi hô hấp thường yếu hơn từ PPG ngón tay so với PPG tai [8,28], điều này có thể giúp giải thích sự thiếu thành công của việc ước tính chu kỳ nhiệm vụ trong trường hợp này.

Điều quan trọng cần lưu ý là các sự khác biệt khác về dạng sóng capnography như mẫu răng cưa, liên quan đến các rối loạn hô hấp tắc nghẽn, và những thay đổi biên độ CO₂ hiếm gặp không được nắm bắt bởi mô hình này. Trong trường hợp các giá trị tuyệt đối cụ thể của nồng độ CO₂ trong hơi thở, hợp lý khi giả định rằng thông tin này sẽ không được mã hóa trong tín hiệu PPG. Tuy nhiên, trong trường hợp dạng sóng răng cưa, có thể thông tin liên quan được chứa trong photoplethysmography, và do đó trong trường hợp này có khả năng không được nắm bắt bởi mô hình do thiếu các ví dụ trong bộ dữ liệu.

Hiện tại, mô hình đã được huấn luyện này sẽ có thể áp dụng trong các kịch bản mà photoplethysmography ngón tay được ghi lại và được lọc thông cao để loại bỏ độ lệch DC. Điều quan trọng cần lưu ý là sẽ cần thêm dữ liệu để huấn luyện mô hình cho các vị trí PPG khác, chẳng hạn như tai, nơi mà dạng sóng PPG có cấu trúc khác và các biến đổi hô hấp chủ đạo hơn hiện diện [8]. Hơn nữa, mô hình hiện tại chỉ được huấn luyện trên dữ liệu mà bệnh nhân chủ yếu đứng yên, và do đó để hiệu quả trong các kịch bản có nhiều nhiễu do chuyển động (như trong các bài kiểm tra đi bộ 6 phút), có khả năng sẽ cần thêm dữ liệu huấn luyện từ các kịch bản này.

Độ phức tạp tương đối thấp của mô hình học sâu được trình bày trong công trình này dẫn đến tốc độ triển khai cực kỳ nhanh. Thời gian trung bình để mô hình tính toán một cửa sổ dạng sóng hô hấp 9,6 giây là 0,43 ms trên bộ xử lý Intel i7-1165G7 (2,8 GHz), và do đó có thể tạo ra hơn 6 giờ dạng sóng hô hấp trong chưa đầy 1 giây. Điều này so với 5 phút dữ liệu mỗi giây khi sử dụng làm mịn Kalman cho các ước tính tần số khác nhau (2,4 GHz) [20] và 4 phút dữ liệu mỗi giây khi sử dụng phương pháp phân rã mode thực nghiệm kết hợp với phân tích thành phần độc lập (3,2 GHz) [29]. Trong thực tế, tốc độ của phương pháp đề xuất của chúng tôi sẽ cho phép triển khai dễ dàng trên các vi điều khiển có kích thước nhỏ trong các thiết bị đeo.



# **Mô tả chi tiết Hình 4: Đánh giá trên bộ dữ liệu BIDMC**

Hình này gồm hai phần chính (a) và (b) trình bày kết quả đánh giá mô hình trên bộ dữ liệu BIDMC:

## **Phần (a): Sai số ước tính tần số hô hấp trên BIDMC**

* Biểu đồ này thể hiện sai số ước tính tần số hô hấp (đơn vị: nhịp thở/phút) theo độ dài cửa sổ thời gian (giây).
* Có 4 đường biểu diễn:
  + Đường liền màu tím: "Train C: Median" - sai số trung vị khi huấn luyện chỉ trên Capnobase
  + Đường đứt màu tím: "Train C: Median MAE" - sai số tuyệt đối trung bình trung vị khi huấn luyện chỉ trên Capnobase
  + Đường liền màu xanh: "Train C+B: Median" - sai số trung vị khi huấn luyện trên cả Capnobase và BIDMC
  + Đường đứt màu xanh: "Train C+B: Median MAE" - sai số tuyệt đối trung bình trung vị khi huấn luyện trên cả Capnobase và BIDMC
* Quan sát thấy sai số giảm đáng kể khi tăng độ dài cửa sổ thời gian, và mô hình huấn luyện trên cả hai bộ dữ liệu (C+B) có hiệu suất tốt hơn so với chỉ huấn luyện trên Capnobase.

## **Phần (b): Ví dụ về dạng sóng từ BIDMC**

Biểu đồ này gồm ba phần nhỏ:

1. Phần trên cùng:
   * Đường đứt màu đen: Dạng sóng từ impedance pneumography (tín hiệu tham chiếu)
   * Đường liền màu đỏ: Đầu ra của mô hình corr-encoder khi chỉ huấn luyện trên Capnobase (Train C)
2. Phần giữa:
   * Đường đứt màu đen: Dạng sóng từ impedance pneumography (tín hiệu tham chiếu)
   * Đường liền màu tím: Đầu ra của mô hình corr-encoder khi huấn luyện trên cả Capnobase và BIDMC (Train C+B)
3. Phần dưới cùng:
   * Đường màu xanh: Tín hiệu đầu vào photoplethysmography (PPG)

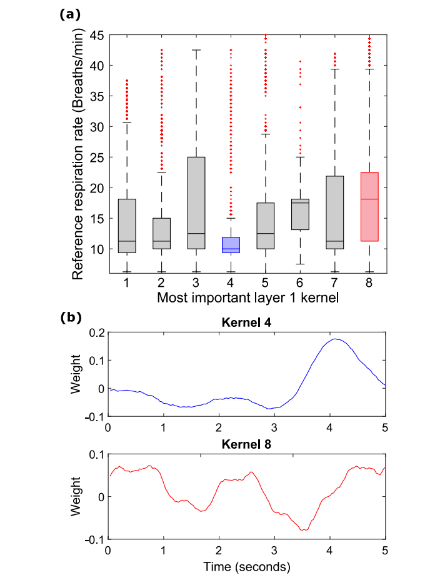
Trục x là thời gian (giây), từ 200 đến 240 giây. Trục y là biên độ tín hiệu (đơn vị tương đối a.u.).

Hình này minh họa rằng ngay cả khi chỉ huấn luyện trên Capnobase (không thấy BIDMC), mô hình vẫn nắm bắt được phần lớn thông tin hô hấp, mặc dù còn một số khác biệt về hình thái sóng. Khi huấn luyện thêm trên BIDMC, mô hình tạo ra dạng sóng gần hơn với impedance pneumography. Điều này chứng minh khả năng tổng quát hóa và thích nghi của mô hình đối với các bộ dữ liệu khác nhau.

# **Hình 4. Kết quả ước tính tần số hô hấp và ví dụ về dạng sóng đầu ra cho bộ dữ liệu BIDMC**

(a) Biểu đồ thể hiện sai số tuyệt đối trung vị trên tất cả các cửa sổ thời gian (đường liền) và sai số tuyệt đối trung bình trung vị trên tất cả các đối tượng (đường đứt). Biểu đồ so sánh giữa hai trường hợp: huấn luyện trên Capnobase và kiểm tra trên BIDMC (màu đỏ), và huấn luyện trên Capnobase sau đó huấn luyện lại trên BIDMC thông qua phương pháp kiểm chứng chéo loại bỏ từng đối tượng (màu tím).

(b) Ví dụ minh họa các dạng sóng đầu ra từ hai phương pháp huấn luyện: huấn luyện chỉ trên Capnobase và kiểm tra trên BIDMC (màu đỏ), và huấn luyện trên Capnobase sau đó huấn luyện lại trên BIDMC thông qua phương pháp kiểm chứng chéo loại bỏ từng đối tượng (màu tím). Các dạng sóng này được so sánh với tiêu chuẩn vàng impedance pneumography (màu xám - đường đứt) và tín hiệu đầu vào photoplethysmography (màu xanh).



# **Hình 5. Khả năng giải thích các trọng số nhân (kernel) theo tần số hô hấp**

(a) Biểu đồ hộp thể hiện các tần số hô hấp tham chiếu mà tại đó một nhân (kernel) cụ thể ở lớp 1 tạo ra giá trị lớn nhất trong không gian ẩn. Cách tính toán này được giải thích chi tiết trong phần phụ "Khả năng giải thích mô hình" của phần Thảo luận.

(b) Trọng số của nhân 4 (màu xanh dương) có ái lực mạnh đối với tần số thấp, và nhân 8 (màu đỏ) có ái lực mạnh đối với tần số hô hấp cao hơn. Các trọng số nhân trong những biểu đồ này đầu tiên được làm mịn bằng bộ lọc trung bình di động có độ dài 1 giây.

# **Mô tả chi tiết Hình 5: Phân tích các nhân (kernel) quan trọng trong mô hình**

## **Phần (a): Biểu đồ hộp về tần số hô hấp tham chiếu**

* Biểu đồ này thể hiện phân bố tần số hô hấp tham chiếu (đơn vị: nhịp thở/phút) liên quan đến 8 nhân quan trọng nhất trong lớp 1 của mô hình neural network.
* Trục y: Tần số hô hấp tham chiếu, dao động từ khoảng 10 đến 45 nhịp thở/phút.
* Trục x: 8 nhân (kernel) quan trọng nhất ở lớp đầu tiên của mạng neural, đánh số từ 1 đến 8.
* Mỗi hộp (box plot) thể hiện:
  + Đường ngang trong hộp: giá trị trung vị
  + Chiều cao của hộp: khoảng tứ phân vị (phạm vi giữa 25% và 75%)
  + Các đường thẳng đứng kéo dài từ hộp: phạm vi giá trị không phải ngoại lệ
  + Các điểm riêng lẻ: các giá trị ngoại lệ
* Đáng chú ý, nhân số 4 (được tô màu xanh dương) có phạm vi hẹp nhất, cho thấy độ ổn định cao nhất trong ước tính tần số hô hấp.

## **Phần (b): Biểu diễn trọng số của các nhân theo thời gian**

* Biểu đồ này thể hiện trọng số của nhân 4 và nhân 8 theo thời gian (0-5 giây).
* Biểu đồ trên (màu xanh dương): "Kernel 4" - Trọng số của nhân 4 theo thời gian, có một đỉnh đặc trưng ở khoảng 3.5-4 giây.
* Biểu đồ dưới (màu đỏ): "Kernel 8" - Trọng số của nhân 8 theo thời gian, có nhiều dao động và một điểm thấp đáng chú ý ở khoảng 3 giây.

Hình này giúp hiểu được cách mạng neural học các mẫu hình đặc trưng trong dữ liệu hô hấp. Nhân 4 (màu xanh) là đặc biệt quan trọng vì nó có sự tương ứng nhất quán nhất với tần số hô hấp tham chiếu, như được minh họa bằng phạm vi hẹp trong biểu đồ hộp. Các mẫu hình trọng số theo thời gian cho thấy những đặc điểm mà mỗi nhân đang phát hiện trong tín hiệu PPG để ước tính thông tin hô hấp.

# **4.1. Khả năng giải thích mô hình**

Đặc biệt đối với các ứng dụng y tế, việc cố gắng làm sáng tỏ "hộp đen" của các phương pháp học sâu là rất quan trọng. Với độ phức tạp tương đối thấp của mô hình này và các nhân (kernel) tích chập có độ dài tương đương với đầu vào, chúng tôi có thể bắt đầu suy luận những gì mô hình đang lựa chọn từ dữ liệu đầu vào. Để tính toán điều này, các trọng số đã được huấn luyện cho 3 lớp đầu tiên của mô hình được trích xuất. Đối với tất cả các cửa sổ đầu vào photoplethysmography từ bộ dữ liệu Capnobase, phép tích chập được thực hiện với tất cả các tổ hợp khác nhau có thể có của các nhân tích chập theo thứ tự. Với 8 nhân khác nhau trong mỗi lớp, điều này dẫn đến 512 phép chiếu khác nhau có thể có của một đầu vào cụ thể lên không gian ẩn.

Trong khi đa số các tổ hợp dẫn đến các giá trị bằng không và không đáng kể, một số đường dẫn chiếu cho một đầu vào nhất định dẫn đến các giá trị ẩn lớn. Giá trị tối đa trên tất cả 512 tổ hợp sau đó được tính toán và liên kết với nhân cụ thể ở lớp 1 đã kích hoạt giá trị tối đa này. Điều này cho phép chúng tôi suy luận các trọng số nhân lớp 1 quan trọng nhất cho một đầu vào nhất định. Các nhân lớp 1 quan trọng nhất sau đó được vẽ biểu đồ so với tần số hô hấp tham chiếu tương ứng cho mỗi đầu vào, dẫn đến các biểu đồ hộp được hiển thị trong Hình 5(a).

Với việc độ dài nhân được chọn có kích thước tương tự như đầu vào, có thể thấy rằng một phần các nhân này đang phân biệt các đầu vào khác nhau dựa trên tần số. Suy luận này được củng cố thêm khi chúng tôi trực quan hóa các trọng số đã được làm mịn của nhân 4 (kích hoạt chủ yếu ở tần số hô hấp thấp với tần số hô hấp tham chiếu trung vị là 10 nhịp thở mỗi phút) và nhân 8 (kích hoạt chủ yếu ở tần số cao hơn với tần số hô hấp tham chiếu trung vị là 18 nhịp thở mỗi phút) trong Hình 5(b).

Các trọng số của nhân 4 tương ứng với mẫu hình tần số thấp của hô hấp, trong khi các trọng số của nhân 8 chứa thông tin tần số cao hơn. Kết quả này quan trọng vì nó cho thấy mô hình đầu vào đang trích xuất một phần thông tin dựa trên tần số từ đầu vào, tương tự như việc thực hiện biến đổi cosin rời rạc. Kiến thức này giúp chúng tôi chuyển từ cách giải thích mô hình "hộp đen" sang "hộp xám".

# **5. Kết luận**

Chúng tôi đã giới thiệu một khuôn khổ học sâu độ phức tạp thấp mới để trích xuất thông tin hô hấp từ photoplethysmography. Mô hình đề xuất đã được chứng minh là tạo ra dạng sóng hô hấp chính xác và các ước tính tần số hô hấp tiên tiến nhất. Chúng tôi đã chứng minh hiệu quả này trên hai bộ dữ liệu khác nhau, Capnobase và BIDMC, làm nổi bật khả năng tổng quát hóa mạnh mẽ của nó.

Quan trọng là, với tính đơn giản tương đối của mô hình, nó đã được chứng minh có khả năng chuyển đổi nhiều giờ dữ liệu photoplethysmography thành dạng sóng hô hấp tương ứng chỉ trong một giây. Khuôn khổ trình bày trong công trình này giúp thu hẹp khoảng cách giữa photoplethysmography và các tham chiếu hô hấp tiêu chuẩn vàng như capnography và impedance pneumography.

Chúng tôi tin rằng mô hình của chúng tôi do đó có thể cải thiện tính hữu ích của các thiết bị đeo dựa trên PPG trong các ứng dụng y tế.

References

[1] M.J. Tobin, T.S. Chadha, G. Jenouri, S.J. Birch, H.B. Gazeroglu, M.A. Sackner,

Breathing patterns: 2. Diseased subjects, Chest 84 (3) (1983) 286–294.

[2] H. Wilkens, B. Weingard, A. Lo Mauro, E. Schena, A. Pedotti, G.W. Sybrecht,

A. Aliverti, Breathing pattern and chest wall volumes during exercise in pa-

tients with cystic fibrosis, pulmonary fibrosis and COPD before and after lung

transplantation, Thorax 65 (9) (2010) 808–814.

[3] D.J. Meredith, D. Clifton, P. Charlton, J. Brooks, C.W. Pugh, L. Tarassenko,

Photoplethysmographic derivation of respiratory rate: A review of relevant

physiology, J. Med. Eng. Technol. 36 (1) (2012) 1–7.

[4] K.H. Shelley, A.A. Awad, R.G. Stout, D.G. Silverman, The use of joint time

frequency analysis to quantify the effect of ventilation on the pulse oximeter

waveform, J. Clin. Monit. Comput. 20 (2) (2006) 81–87.

[5] L. Nilsson, A. Johansson, S. Kalman, Monitoring of respiratory rate in postopera-

tive care using a new photoplethysmographic technique, J. Clin. Monit. Comput.

16 (4) (2000) 309–315.

[6] N. Rehman, D.P. Mandic, Multivariate empirical mode decomposition, Proc. R.

Soc. A 466 (2117) (2010) 1291–1302.

[7] M.A. Motin, C.K. Karmakar, M. Palaniswami, Selection of empirical mode

decomposition techniques for extracting breathing rate from PPG, IEEE Signal

Process. Lett. 26 (4) (2019) 592–596.

[8] H.J. Davies, P. Bachtiger, I. Williams, P.L. Molyneaux, N.S. Peters, D.P. Mandic,

Wearable in-ear PPG: Detailed respiratory variations enable classification of

COPD, IEEE Trans. Biomed. Eng. 69 (7) (2022) 2390–2400.

[9] P.H. Charlton, D.A. Birrenkott, T. Bonnici, M.A.F. Pimentel, A.E.W. Johnson, J.

Alastruey, L. Tarassenko, P.J. Watkinson, R. Beale, D.A. Clifton, Breathing rate

estimation from the electrocardiogram and photoplethysmogram: A review, IEEE

Rev. Biomed. Eng. 11 (2018) 2–20.

[10] M.S. Siobal, Monitoring exhaled carbon dioxide, Respir. Care 61 (10) (2016)

1397–1416.

[11] H. Aminiahidashti, S. Shafiee, A. Zamani Kiasari, M. Sazgar, Applications of end-

tidal carbon dioxide (ETCO2) monitoring in emergency department: A narrative

review, Emerg. (Tehran) 6 (1) (2018) e5.

[12] A. Gogna, A. Majumdar, R. Ward, Semi-supervised stacked label consistent

autoencoder for reconstruction and analysis of biomedical signals, IEEE Trans.

Biomed. Eng. 64 (9) (2017) 2196–2205.

[13] D.-H. Kang, D.-H. Kim, 1D convolutional autoencoder-based PPG and GSR signals

for real-time emotion classification, IEEE Access 10 (2022) 91332–91345.

[14] L. Gondara, Medical image denoising using convolutional denoising autoen-

coders, in: Proc of the 16th IEEE International Conference on Data Mining

Workshops (ICDMW), 2016, pp. 241–246.

[15] H.-T. Chiang, Y.-Y. Hsieh, S.-W. Fu, K.-H. Hung, Y. Tsao, S.-Y. Chien, Noise

reduction in ECG signals using fully convolutional denoising autoencoders, IEEE

Access 7 (2019) 60806–60813.

[16] W. Karlen, S. Raman, J.M. Ansermino, G.A. Dumont, Multiparameter respiratory

rate estimation from the photoplethysmogram, IEEE Trans. Biomed. Eng. 60 (7)

(2013) 1946–1953.

[17] A. Johansson, P.A. Öberg, Estimation of respiratory volumes from the photo-

plethysmographic signal. Part 2: A model study, Med. Biol. Eng. Comput. 37 (1)

(1999) 48–53.

[18] R. Katiyar, V. Gupta, R.B. Pachori, FBSE-EWT-based approach for the deter-

mination of respiratory rate from PPG signals, IEEE Sens. Lett. 3 (7) (2019)

1–4.

[19] M.A. Motin, C.K. Karmakar, M. Palaniswami, Ensemble empirical mode decom-

position with principal component analysis: A novel approach for extracting

respiratory rate and heart rate from photoplethysmographic signal, IEEE J.

Biomed. Health Inf. 22 (3) (2018) 766–774.

[20] M.A.F. Pimentel, A.E.W. Johnson, P.H. Charlton, D. Birrenkott, P.J. Watkinson,

L. Tarassenko, D.A. Clifton, Toward a robust estimation of respiratory rate from

pulse oximeters, IEEE Trans. Biomed. Eng. 64 (8) (2016) 1914–1923.

[21] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark,

J.E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, PhysioBank, PhysioToolkit,

and PhysioNet, Circulation 101 (23) (2000) e215–e220.

[22] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin,

N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison,

A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, PyTorch:

An imperative style, high-performance deep learning library, in: Advances in

Neural Information Processing Systems 32, Curran Associates, Inc., 2019, pp.

8024–8035.

[23] S. de Jong, SIMPLS: An alternative approach to partial least squares regression,

Chemometr. Intell. Lab. Syst. 18 (3) (1993) 251–263.

[24] S. Khreis, D. Ge, H.A. Rahman, G. Carrault, Breathing rate estimation using

Kalman smoother with electrocardiogram and photoplethysmogram, IEEE Trans.

Biomed. Eng. 67 (3) (2019) 893–904.

[25] S.A.H. Aqajari, R. Cao, A.H.A. Zargari, A.M. Rahmani, An end-to-end and

accurate PPG-based respiratory rate estimation approach using cycle generative

adversarial networks, in: Proceedings of the 43rd Annual International Confer-

ence of the IEEE Engineering in Medicine and Biology Society, EMBC, 2021, pp.

744–747.

[26] J. Johnson, A. Alahi, L. Fei-Fei, Perceptual losses for real-time style transfer

and super-resolution, in: Computer Vision – ECCV 2016, Springer International

Publishing, 2016, pp. 694–711.

[27] P. Isola, J.-Y. Zhu, T. Zhou, A.A. Efros, Image-to-image translation with condi-

tional adversarial networks, in: Proceedings of the IEEE Conference on Computer

Vision and Pattern Recognition, CVPR, 2017.

[28] K. Budidha, P.A. Kyriacou, In vivo investigation of ear canal pulse oximetry

during hypothermia, J. Clin. Monit. Comput. 32 (1) (2018) 97–107.

[29] R. Lei, B.W.-K. Ling, P. Feng, J. Chen, Estimation of heart rate and respiratory

rate from PPG signal using complementary ensemble empirical mode decom-

position with both independent component analysis and non-negative matrix

factorization, Sensors 20 (11) (2020) 3238.